

Seturi de date

Una dintre principalele aspecte care trebuie solutionate si avute in vedere in alegerea unei teme si in realizarea ei cu succes este identificarea (daca exista) a unui set de date.

Pentru anumite subdomenii nu exista seturi de date (sau cel putin not for free) astfel incat, in anumite contexte, colectarea seturilor de date este un task esential in sine care aduce plus valoare proiectului.

Exista cateva site-uri care permit cautarea datelor folosind anumite criterii, cum ar fi anumite cuvinte cheie.

Denumire	Adresa
Google datasets	https://datasetsearch.research.google.com/
Data world	https://data.world/
Data hub	https://datahub.io/
Climate Knowledge Portal	https://climateknowledgeportal.worldbank.org/download-data
Kaggle Datasets	https://www.kaggle.com/datasets
UCI Repositories	https://archive.ics.uci.edu/ml/datasets.php

Pentru accesarea si downloadul seturilor de date pe unele platforme va trebui sa va logati intai (recomandarea ar fi sa va logati folosind adresa de email de facultate).

O lista mai extinsa (pe multiple domenii) poate sa fie accesata si aici: ***Top 75 public datasets for machine learning*** - [https://blog.superannotate.com/public-datasets-for-machine-learning/?utm_term=&utm_campaign=Annotation Search New Segmented&utm_source=adwords&utm_medium=ppc&hsa_a_cc=7527629942&hsa_cam=14073003037&hsa_grp=132203422229&hsa_ad=577732942220&hsa_src=g&hsa_tgt=dsa-392284169515&hsa_kw=&hsa_mt=&hsa_net=adwords&hsa_ver=3&gclid=Cj0KCQiA09eQBhCxARIsAAYRiymkflbyWIOUiVLXpLvoGfcYRjIrcnyVtoKeXhD6p1PPv1OSwihIOBYaAiYgEALw_wcB](https://blog.superannotate.com/public-datasets-for-machine-learning/?utm_term=&utm_campaign=Annotation+Search+New+Segmented&utm_source=adwords&utm_medium=ppc&hsa_a_cc=7527629942&hsa_cam=14073003037&hsa_grp=132203422229&hsa_ad=577732942220&hsa_src=g&hsa_tgt=dsa-392284169515&hsa_kw=&hsa_mt=&hsa_net=adwords&hsa_ver=3&gclid=Cj0KCQiA09eQBhCxARIsAAYRiymkflbyWIOUiVLXpLvoGfcYRjIrcnyVtoKeXhD6p1PPv1OSwihIOBYaAiYgEALw_wcB)